Virtual Full Duplex Wireless Broadcast via Sparse Recovery

Lei Zhang and Dongning Guo
Department of Electrical Engineering & Computer Science
Northwestern University
Evanston, IL 60208, USA

***Abstract*—A novel solution is proposed to achieve a frequent task in wireless networks, where all node broadcast information to and receive information from nodes within a single hop. The solution exploits the multiaccess nature of the wireless medium and addresses the half-duplex constraint at the fundamental level. The defining feature of the scheme is to let all nodes send their messages at the same time, where each node broadcasts a codeword (selected from its unique codebook) consisting of on-slots and off-slots. A node transmits only during its on-slots, and listens to its peers through its own off-slots. Each node decodes the messages of its peers based on the superposed signals received through its own off-slots. Decoding can be viewed as a problem of *sparse support recovery* based on linear measurements. In the case that each message consists of a small number of bits, an iterative message-passing algorithm based on belief propagation is developed. In a network consisting of Poisson distributed nodes, numerical results demonstrate that the proposed scheme achieves several times the rate of slotted ALOHA and CSMA with the same packet error rate (1%).**

I. Introduction

Consider a frequent situation in wireless peer-to-peer networks, where every node wishes to broadcast messages to all nodes within its one-hop neighborhood, called its *peers*, and also wishes to receive messages from its peers. We refer to this problem as *mutual broadcast*. Such traffic can be dominant in many applications, such as messaging or video conferencing of multiple parties in a spontaneous social network, or on an incident scene or a battlefield. Wireless mutual broadcast is also critical to efficient network resource allocation, where messages are exchanged between nodes about their demands and local states, such as queue length, channel quality, code and modulation format, and request for certain resources and services.

A major challenge in wireless networks is the half-duplex constraint, namely, currently affordable radio cannot receive useful signals at the same time over the same frequency band over which it is transmitting. This is largely due to the limited dynamic range and noise of the radio frequency circuits, which are likely to remain a physical restriction in the near future. An important consequence of the half-duplex constraint in a typical implementation of wireless networks is that, if two peers transmit their packets at the same time, they do not hear each other. To achieve reliable mutual broadcast using a usual packet-based scheme, nodes have to repeat their packets many times interleaved with random delays, so that all peers can hear each other after enough retransmissions. This is basically the ubiquitous random channel access solution.

A closer examination of the half-duplex constraint, however, reveals that a node does not need to transmit an entire packet before listening to the channel. An alternative solution is conceivable: Let a frame (typically of a few thousand symbols) be divided into some number of slots, where each node transmits over a subset of the slots and assumes silence over the remaining slots, then the node can receive useful signals over those nontransmission slots. If nodes activate different sets of on-slots, then nodes can all transmit information during a frame and receive useful signals within the same frame, and decode messages from peers as long as sufficiently strong error-control codes are applied. This on-off signaling, called *rapid on-off-division duplex (RODD)*, was originally proposed in [2]. Using RODD, reliable mutual broadcast can be achieved using a single frame interval.

Importantly, RODD achieves virtual full-duplex communication using half-duplex radios. Despite the half-duplex physical layer, the radio appears to be full-duplex in higher layers. Not only is RODD signaling applicable to the mutual broadcast problem, it can also be the basis of a clean-slate design of the physical and medium access control (MAC) layers of wireless peer-to-peer networks. RODD-based schemes have advantages over state-of-the-art designs of MAC protocols, which either apply ALOHA-type random access or use a mixture of random access and scheduling/reservation. This is in part because RODD eliminates retransmissions due to collisions in random access schemes.

In order to decode messages from neighbors, it is necessary to acquire their timing (or relative delay) regardless of whether RODD or any other physical- and MAC-layer technology is used. Timing acquisition and decoding are generally easier if the frames arriving at a receiver are synchronous locally within each neighborhood, although synchronization is not a necessity. In a wireless network, synchronization can be achieved using various distributed algorithms for reaching consensus [3]–[5] or using a common source of timing, such as the Global Positioning System (GPS). Whether synchronizing the nodes is worthwhile is a challenging question, which is not discussed further in this paper.

In this paper, we focus on a special use of RODD signaling and a special case of mutual broadcast, where each node has a small number of bits to send to its peers. It is assumed that

This work has been presented in part at the 2011 IEEE International Symposium on Information Theory. [1]

node transmissions are perfectly synchronized. The goal here is to provide a practical algorithm for encoding and decoding the short messages to achieve reliable and efficient mutual broadcast. Decoding is in fact a problem of support recovery based on linear measurements, since the received signal is basically a noisy superposition of peers' codewords selected from their respective codebooks. There are many algorithms developed in the compressed sensing (or *sparse recovery*) literature to solve the problem, the complexity of which is often polynomial in the size of the codebook (see, e.g., [6]–[10]). In this paper, an iterative message-passing algorithm based on belief propagation (BP) with linear complexity is developed. Numerical results show that the proposed RODD scheme significantly outperforms slotted-ALOHA with multi-packet reception capability and CSMA in terms of data rate.

The excellent performance of the proposed scheme is because it departs from the usual solution where a highly reliable, highly redundant, capacity-achieving, point-to-point physical-layer code is paired with a rather unreliable MAC layer. By treating the physical and MAC layers as a whole, the proposed scheme achieves better overall reliability at much higher efficiency.

The remainder of the paper is organized as follows. After the system model is presented in Section II, Section III studies the conventional random-access schemes, namely slotted-ALOHA with multi-packet reception capability and CSMA. In Section IV, the proposed coding scheme for mutual broadcast is described. The message-passing decoding algorithm is developed in Section V. Numerical comparisons are presented in Section VI. Section VII concludes the paper.

II. Channel and Network Models

A. Linear Channel Model

Let $\Phi = \{X_i\}_i$ denote the set of nodes on the plane. We refer to a node by its location X_i. Suppose all transmissions use the same single carrier frequency. Let time be slotted and all nodes be perfectly synchronized.[1] Let $\varpi_i \in \{1, \dots, 2^l\}$ denote the message node X_i wishes to broadcast. In discrete-time baseband, let $\boldsymbol{S}_i(\varpi_i)$ denote the signature (codeword) transmitted by node X_i, whose entries take values in $\{-1, 0, +1\}$. Let U_{0i} denote the complex-valued coefficient of the wireless link from X_i to X_0. The signal received by node X_0, if it could listen over the entire frame, is then described by

$$\widetilde{\boldsymbol{Y}} = \sqrt{\gamma} \sum_{X_i \in \Phi \setminus \{X_0\}} U_{0i} \boldsymbol{S}_i(\varpi_i) + \widetilde{\boldsymbol{W}} \tag{1}$$

where $\widetilde{\boldsymbol{W}}$ is noise consisting of independent identically distributed (i.i.d.) circularly symmetric complex Gaussian entries with zero mean and unit variance, and γ denotes the nominal signal-to-noise ratio (SNR). Denote the set of neighbors of X_0 by $\mathcal{N}(X_0)$. For simplicity, if we further assume that transmissions from non-neighbors, if any, are accounted for as part of the additive Gaussian noise, (1) can be rewritten as

$$\widetilde{\boldsymbol{Y}} = \sqrt{\gamma} \sum_{X_i \in \mathcal{N}(X_0)} U_{0i} \boldsymbol{S}_i(\varpi_i) + \overline{\boldsymbol{W}} \tag{2}$$

[1]See [2] for a discussion of synchronization issues. In [11], cyclic codes are proposed to resolve the user delays in a multiaccess channel.

where each element in $\overline{\boldsymbol{W}}$ is assumed to be circularly symmetric complex Gaussian with variance σ^2. The variance accounts for interference from non-neighbors and depends on the network topology. It will be derived in Section IV.

B. Network Model

Consider a network with nodes distributed across the plane according to a homogeneous Poisson point process (p.p.p.) with intensity λ. The number of nodes in any region of area A is a Poisson random variable with mean λA. Without loss of generality, we assume node X_0 is located at the origin and focus on its performance, which should be representative of any node in the network.

Poisson point process is the most frequently used model to study wireless networks (see [12] and references therein). The RODD signaling and the mutual broadcast scheme proposed in this paper are of course not limited to homogeneous Poisson distributed networks. The homogeneous p.p.p. model is assumed in this paper to facilitate analysis and comparison of RODD and competing technologies.

C. Propagation Model and Neighborhood

It is assumed that the large-scale signal attenuation over distance follows power law with some path-loss exponent $\alpha > 2$, and the small-scale fading of a link is modeled by a Rayleigh random variable with mean equal to 1. There are different ways to define the neighborhood of a node. For concreteness, we say that nodes X_i and X_j are neighbors of each other if the channel gain between them is no less than a certain threshold, denoted by θ. Link reciprocity is regarded as given.

For any pair of nodes $X_i, X_j \in \Phi$, let $R_{ij} = |X_i - X_j|$ and G_{ij} denote the distance and the small-fading gain between them in a given frame, respectively. Then the channel gain from X_j to X_i is $G_{ij} R_{ij}^{-\alpha}$. The neighborhood of a node depends on the instantaneous fading gains. Specifically, we denote the set of neighbors of node X_i as

$$\mathcal{N}(X_i) = \left\{ X_j \in \Phi : G_{ij} R_{ij}^{-\alpha} \geq \theta, j \neq i \right\}. \tag{3}$$

The channel coefficient U_{ij} should satisfy $|U_{ij}|^2 = G_{ij} R_{ij}^{-\alpha}$, where its phase is assumed to be uniformly distributed on $[0, 2\pi)$ independent of everything else. Assuming the Poisson point process network model introduced in Section II-B, the distribution of the amplitude of channel coefficient U_{0i} in (1) for an arbitrary neighbor $X_i \in \mathcal{N}(X_0)$ is derived in the following.

Without loss of generality, we drop the indices 0 and i, and use R and G to denote the distance and the fading gain, respectively. Since the two nodes are assumed to be neighbors, G and R satisfy $GR^{-\alpha} \geq \theta$, i.e., $R \leq (G/\theta)^{1/\alpha}$. Under the assumption that all nodes form a p.p.p., for given G, this arbitrary neighbor X_i is uniformly distributed in a disc centered at node X_0 with radius $(G/\theta)^{1/\alpha}$. Therefore, the conditional distribution of R given G can be expressed as

$$\mathsf{P}(R \leq r | G) = \min \left\{ 1, r^2 \left(\frac{\theta}{G} \right)^{\frac{2}{\alpha}} \right\}. \tag{4}$$

Now for every $u \geq \sqrt{\theta}$, by (4) we have

$$\begin{aligned} \mathsf{P}(GR^{-\alpha} \geq u^2) &= \mathsf{E}_G \left\{ \mathsf{P}\left(R \leq \left(\frac{G}{u^2}\right)^{\frac{1}{\alpha}} \middle| G \right) \right\} \\ &= \mathsf{E}_G \left\{ \left(\frac{G}{u^2}\right)^{\frac{2}{\alpha}} \left(\frac{\theta}{G}\right)^{\frac{2}{\alpha}} \right\} \\ &= \frac{\theta^{\frac{2}{\alpha}}}{u^{\frac{4}{\alpha}}} . \end{aligned} \tag{5}$$

Hence the probability density function (pdf) of $|U_{0i}|$ of neighbor X_i is

$$p(u) = \begin{cases} \frac{4}{\alpha} \frac{\theta^{2/\alpha}}{u^{4/\alpha+1}}, & u \geq \sqrt{\theta}; \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

In fact, coefficient vector $\mathcal{G}_i = (G_{ji})_j$ for all $j \neq i$ can be regarded as a mark of node X_i, so that $\tilde{\Phi} = \{(X_i, \mathcal{G}_i)\}_i$ is a marked p.p.p. Denote

$$\hat{\Phi} = \tilde{\Phi} \backslash (X_0, \mathcal{G}_0) \tag{7}$$

given that $(X_0, \mathcal{G}_0)$ is at the origin. By the Slivnyak-Meche theorem [12], $\hat{\Phi}$ is also a marked p.p.p. with intensity λ. By the Campbell's theorem [12], the average number of neighbors of X_0 can be obtained as:

$$\begin{aligned} c &= \mathsf{E}_{\hat{\Phi}} \left\{ \sum_{(X_i, \mathcal{G}_i) \in \hat{\Phi}} \mathbf{1}\left(G_{0i} R_{0i}^{-\alpha} \geq \theta\right) \right\} \\ &= 2\pi\lambda \int_0^\infty \int_0^\infty \mathbf{1}\left(g r^{-\alpha} \geq \theta\right) r e^{-g} \mathrm{d}r \mathrm{d}g \\ &= \frac{2}{\alpha} \pi \lambda \theta^{-2/\alpha} \Gamma\left(\frac{2}{\alpha}\right) \end{aligned} \tag{8}$$

where $\mathbf{1}(\cdot)$ is the indicator function and $\Gamma(\cdot)$ is the Gamma function.

III. Random-Access Schemes

In this section we describe two random access schemes, namely slotted ALOHA and CSMA, and provide lower bounds on the error probability for a given number of symbol transmissions. The results will be used in Section VI to compare with the performance of our proposed sparse recovery scheme.

Let L denote the total number of bits encoded into a frame, which includes an l-bit message and a few additional bits which identify the sender. This is in contrast to broadcast via sparse recovery, where the signature itself identifies the sender (and carries the message). Each broadcast period consists of a number of frames to allow for retransmissions. Without loss of generality, we still consider the typical node X_0 at the origin. An error event is defined as that node X_0 cannot correctly recover the message from one specific neighbor. The corresponding error probabilities achieved by slotted ALOHA and CSMA are denote by $\mathsf{P}_\mathrm{a}^\mathrm{e}$ and $\mathsf{P}_\mathrm{c}^\mathrm{e}$, respectively.

A. Slotted ALOHA

In slotted ALOHA, suppose each node chooses independently with the same probability p to transmit in every frame interval. A message is assumed to be decoded correctly if the signal-to-interference-plus-noise ratio (SINR) in the corresponding frame transmission is no smaller than a threshold δ (multi-packet reception is possible only if $\delta < 1$). Over the additive white noise channel with SINR δ, in order to send L bits reliably through the channel, the number of symbols in a frame must exceed

$$\frac{L}{\log_2(1+\delta)} . \tag{9}$$

Let X denote one specific neighbor of node X_0 and G denote the fading coefficient between them. Suppose the mark of X is denoted by $\mathcal{G}$. Given that $(X, \mathcal{G}) \in \hat{\Phi}$ where $\hat{\Phi}$ is given by (7), denote $\hat{\Phi}_1 = \hat{\Phi} \backslash \{(X, \mathcal{G})\}$, which is also a marked p.p.p. with intensity λ. For a given realization of (X, G) and $\hat{\Phi}_1$, define $\mathsf{P}_\mathrm{a}^\mathrm{s}(X, G, \hat{\Phi}_1)$ as the probability that the received SINR from X to X_0 is no less than the threshold δ conditioning on that X transmits in a given frame. In any given frame, the probability of the event that X transmits, X_0 listens, and the transmission is successful is thus $p(1-p)\mathsf{P}_\mathrm{a}^\mathrm{s}(X, G, \hat{\Phi}_1)$. Therefore, the probability that the message from X has not been successfully received by X_0 after M_f consecutive frame intervals can be expressed as

$$\mathsf{P}_\mathrm{a}^\mathrm{e} = \mathsf{E}\left\{ \left(1 - p(1-p)\mathsf{P}_\mathrm{a}^\mathrm{s}(X, G, \hat{\Phi}_1)\right)^{M_f} \right\} \tag{10}$$

where the expectation is over the joint distribution $(X, G, \hat{\Phi}_1)$. Due to the convexity of function $(\max\{0, 1-z\})^n$, $z \geq 0, n \in \{1, 2, \dots\}$, $\mathsf{P}_\mathrm{a}^\mathrm{e}$ in (10) can be lower bounded as

$$\mathsf{P}_\mathrm{a}^\mathrm{e} \geq \left(\max\left\{0, 1 - p(1-p)\mathsf{E}\left\{\mathsf{P}_\mathrm{a}^\mathrm{s}(X, G, \hat{\Phi}_1)\right\}\right\}\right)^{M_f} . \tag{11}$$

In Appendix A, the expectation of $\mathsf{P}_\mathrm{a}^\mathrm{s}(X, G, \hat{\Phi}_1)$ is calculated using the known Laplace transform of the distribution of the interference [12]. For a given number of symbol transmissions M_a, the lower bound on $\mathsf{P}_\mathrm{a}^\mathrm{e}$ is presented in the following result.

Proposition 1: Consider an arbitrary neighbor X of node X_0. The probability that X_0 cannot successfully receive the message from X after M_a symbol transmissions is lower bounded as follows:

$$\begin{aligned} \mathsf{P}_\mathrm{a}^\mathrm{e} \geq \Bigg(\max\Bigg\{ 0, 1 - \frac{1}{\pi} p(1-p) \left(\frac{\theta}{\delta}\right)^b \sin\left(\frac{b\pi}{2}\right) \Gamma(1-b) \\ \int_{-\infty}^{\infty} |\omega|^{b-1} \exp\left\{ -\lambda p \frac{b\pi^2}{\sin(b\pi)} (\iota\omega)^b - \iota \frac{\omega}{\gamma} \right\} \mathrm{d}\omega \Bigg\} \Bigg)^{n_a} \end{aligned} \tag{12}$$

where $\iota = \sqrt{-1}$, $b = 2/\alpha$ and $n_a = M_a \log_2(1+\delta)/L$.

Although (12) appears to be complicated, computing it only involves a straightforward single-variable integral (the outcome of the integral is in fact real-valued).

In the slotted ALOHA scheme, despite repeated transmissions, a given link may still fail to deliver the message due to the half-duplex constraint (the receiver happens to transmit

during the same frame) and consistently weak received SINR due to random interference from other links.

B. CSMA

As an improvement over ALOHA, CSMA lets nodes use a brief contention period to negotiate a schedule in such a way that nodes in a small neighborhood do not transmit data simultaneously. We analyze the performance of CSMA by using the Matérn hard core model [12]. To be specific, consider the following generic scheme: Each node senses the channel continuously; if the channel is busy, the node remains silent and disables its timer; as soon as the channel becomes available, the node starts its timer with a random offset, and waits till the timer expires to transmit its frame. Clearly, the node whose timer expires first in its neighborhood captures the channel and transmits its frame.

Mathematically, let $\{T_i\}$ be i.i.d. random variables with uniform distribution on $[0,1]$, which represent the timer offsets for all nodes $\{X_i\}$ in Φ, respectively. By viewing T_i as a mark of node X_i we redefine $\tilde{\Phi} = \{(X_i, \mathcal{G}_i, T_i)\}_i$, which is still a marked p.p.p. with intensity λ. The medium access indicators $\{e_i\}_i$ are additional dependent marks of the nodes in Φ defined as follows:

$$E_i = \mathbf{1}\left(T_j > T_i,\ \forall X_j \in \mathcal{N}(X_i)\right). \tag{13}$$

The probability of $T_j = T_i$ for $j \neq i$ is zero. Node X_i will transmit its frame if and only if $E_i = 1$.

The same as in the slotted ALOHA case, a message is assumed to be decoded correctly if SINR in the corresponding frame transmission is no smaller than δ. It follows that the number of symbols in a frame must exceed (9).

Let X be one specific neighbor of X_0. Define G and $\mathcal{G}$ as in Section III-A. Given that $(X, \mathcal{G}) \in \hat{\Phi}$, denote $\hat{\Phi}_1 = \hat{\Phi} \backslash \{(X, \mathcal{G})\}$. For a given realization of (X, G) and $\hat{\Phi}_1$, define $\mathsf{P}^{\mathsf{s}}_{\mathsf{c}}(X, G, \hat{\Phi}_1)$ as the probability that node X transmits its frame and the received SINR from X is no less than the threshold δ. Therefore, the probability that the message from X has not been successfully received after M_f consecutive frame intervals can be expressed as

$$\begin{aligned}\mathsf{P}^{\mathsf{e}}_{\mathsf{c}} &= \mathsf{E}\left\{\left(1 - \mathsf{P}^{\mathsf{s}}_{\mathsf{c}}(X, G, \hat{\Phi}_1)\right)^{M_f}\right\} \\ &\geq \left(\max\left\{0, 1 - \mathsf{E}\left\{\mathsf{P}^{\mathsf{s}}_{\mathsf{c}}(X, G, \hat{\Phi}_1)\right\}\right\}\right)^{M_f}. \end{aligned} \tag{14}$$

where the expectation is over the joint distribution of $(X, G, \hat{\Phi}_1)$, and (14) is due to the convexity of function $(\max\{0, 1-z\})^n$, $z \geq 0, n = 0, 1, \ldots$.

For any given number of symbol transmissions M_c, the lower bound on error probability $\mathsf{P}^{\mathsf{e}}_{\mathsf{c}}$ is given by the following result, which is proved in Appendix B.

Proposition 2: Consider an arbitrary neighbor X of node X_0. The probability that X_0 cannot successfully receive the message from X after M_c symbol transmissions is lower bounded as follows:

$$\mathsf{P}^{\mathsf{e}}_{\mathsf{c}} \geq \left(\max\left\{0, 1 - \frac{1}{c^2}\left(\frac{\theta\gamma}{\delta}\right)^{\frac{2}{\alpha}}\left(e^{-c} + c - 1\right)\right\}\right)^{n_c} \tag{15}$$

where c is defined in (8) and $n_c = M_c \log_2(1+\delta)/L$.

In contrast to slotted ALOHA, frame loss due to the half-duplex constraint is eliminated through contention. However, a given link may still fail to deliver the message after repeated transmissions because the received SINR is constantly weak due to random interference outside the neighborhood.

IV. Encoding for Mutual Broadcast

In contrast to random-access schemes, where many retransmissions are needed to achieve a desired error performance, we next describe a unique signaling that allow all message exchanges to finish within one (longer) frame of transmission. The key idea is that each node broadcasts a codeword consisting of on-slots and off-slots. A node transmits only during its on-slots, and listens to its peers through its own off-slots.

Suppose each node X_i is assigned a unique codebook of 2^l on-off signatures (codewords) of length M_s, denoted by $\{\boldsymbol{S}_i(1), \ldots, \boldsymbol{S}_i(2^l)\}$. The optimal design of the signatures is out of the scope of this paper. For simplicity, let each element of each signature be generated randomly and independently, which is 0 with probability $1-q$ and 1 and -1 with probability $q/2$ each. Node X_i broadcasts its l-bit message (or information index) $\varpi_i \in \{1, \ldots, 2^l\}$ by transmitting the codeword $\boldsymbol{S}_i(\varpi_i)$. As discussed in Sections I and II, it is reasonable to assume that node transmissions are synchronized.

In each symbol slot, those transmitting nodes in $\hat{\Phi}$ form an independent thinning of $\hat{\Phi}$ with retention probability q, denoted by $\hat{\Phi}_q$. $\hat{\Phi}_q$ is still an independent marked p.p.p. but with intensity λq. Thus, the sum power from all transmitting non-neighbors of node X_0 in each time slot is derived as

$$\begin{aligned}&\mathsf{E}_{\hat{\Phi}_q}\left\{\sum_{(X_i, \mathcal{G}_i) \in \hat{\Phi}_q} \gamma G_{0i} R_{0i}^{-\alpha} \mathbf{1}\left(G_{0i} R_{0i}^{-\alpha} < \theta\right)\right\} \\ &= 2\pi\lambda q\gamma \int_0^\infty \int_0^\infty g r^{-\alpha} \mathbf{1}\left(g r^{-\alpha} < \theta\right) r e^{-g} \mathrm{d}r \mathrm{d}g \\ &= 2\pi\lambda q\gamma \int_0^\infty r^{-\alpha+1}\left[1 - (\theta r^\alpha + 1) e^{-\theta r^\alpha}\right] \mathrm{d}r \\ &= \frac{4\pi\lambda q\gamma\theta}{\alpha - 2} \int_0^\infty r e^{-\theta r^\alpha} \mathrm{d}r \\ &= \frac{4}{\alpha(\alpha-2)} \pi\lambda q\gamma\theta^{1-2/\alpha} \Gamma\left(\frac{2}{\alpha}\right). \end{aligned} \tag{16}$$

Therefore, the variance of each element of $\overline{\boldsymbol{W}}$ in (2) is

$$\sigma^2 = \frac{4}{\alpha(\alpha-2)} \pi\lambda q\gamma\theta^{1-2/\alpha} \Gamma\left(\frac{2}{\alpha}\right) + 1. \tag{17}$$

The signal received by the typical node X_0, if it could listen over the entire frame, is described by (2). Suppose $|\mathcal{N}(X_0)| = K$ and the neighbors of X_0 are indexed by $1, 2, \ldots, K$. The total number of signatures of all neighbors is $N = 2^l K$. Due to the half-duplex constraint, however, node X_0 can only listen during its off-slots, the number of which has binomial distribution, denoted by $M \sim \mathcal{B}(M_s, 1-q)$, whose expected value is $\mathsf{E}\{M\} = M_s(1-q)$. Let the matrix $\boldsymbol{S} \in \mathbb{R}^{M \times N}$ consist of columns of the signatures from all neighbors of node X_0, observable during the M off-slots of node X_0, and then normalized by $\sqrt{M_s q(1-q)}$ so that the

expected value of the l_2 norm of each column in $\boldsymbol{S}$ is equal to 1. Based on (2), the M-vector observed through all off-slots of node X_0 can be expressed as

$$\boldsymbol{Y} = \sqrt{\gamma_s}\boldsymbol{S}\boldsymbol{X} + \boldsymbol{W} \tag{18}$$

where

$$\gamma_s = \gamma M_s q(1-q)/\sigma^2, \tag{19}$$

and $\boldsymbol{X}$ is an N-vector indicating which K signatures are selected to form the sum in (2) as well as the signal strength for each neighbor. Precisely, $X_{(j-1)2^l+i} = U_j \mathbf{1}(w_j = i)$ for $1 \le j \le K$ and $1 \le i \le 2^l$. For example, consider $K = 3$ neighbors with $l = 2$ bits of information each, where $w_1 = 3, w_2 = 2, w_3 = 1$, then vector $\boldsymbol{X}$ is expressed as

$$\boldsymbol{X} = [0\ \ 0\ \ U_1\ \ 0\ \ 0\ \ U_2\ \ 0\ \ 0\ \ U_3\ \ 0\ \ 0\ \ 0]. \tag{20}$$

The sparsity of $\boldsymbol{X}$ is exactly 2^{-l}, which is very small for large l. The average system load is defined as $\beta = \mathsf{E}\{N\}/\big(M_s(1-q)\big) = 2^l c/\big(M_s(1-q)\big)$.

In general, the decoding problem node X_0 faces is to identify, out of a total of $N = 2^l K$ signatures from all its neighbors, which K signatures were selected. This requires every node to know the codebooks of all neighbors. One solution is to let the codebook of each node be generated using a pseudo-random number generator using its network interface address (NIA) as the seed, so that it suffices to acquire all neighbors' NIAs. This, in turn, is a neighbor discovery problem, which has been studied in [13]–[15]. The discovery scheme proposed in [14], [15] uses similar on-off signalling and also solves a compressed sensing problem.

V. Sparse Recovery (Decoding) via Message Passing

The problem of recovering the support of the sparse input $\boldsymbol{X}$ based on the observation $\boldsymbol{Y}$ has been intensively studied in the compressed sensing literature. In this section, we develop an iterative message-passing algorithm based on belief propagation. The reasons for the choice include: 1) It is one of the most competitive decoding schemes in terms of error performance; and 2) the complexity is only linear in the vector to be estimated.

A. The Factor Graph

Belief propagation belongs to a general class of message-passing algorithms for statistical inference on graphical models, which has demonstrated empirical success in many applications including error-control codes, neural networks, and multiuser detection in code-division multiple access (CDMA) systems.

In order to apply BP to the coded mutual broadcast problem, we construct a Forney-style bipartite factor graph to represent the model (18). Here, we separate the real and imaginary parts in (18) as

$$\boldsymbol{Y}^{(1)} = \sqrt{\gamma_s}\boldsymbol{S}\boldsymbol{X}^{(1)} + \boldsymbol{W}^{(1)},\ \boldsymbol{Y}^{(2)} = \sqrt{\gamma_s}\boldsymbol{S}\boldsymbol{X}^{(2)} + \boldsymbol{W}^{(2)} \tag{21}$$

where the superscripts (1) and (2) represent the real and imaginary parts respectively, $\boldsymbol{W}^{(i)}, i = 1, 2$ consists of i.i.d. Gaussian random variables with zero mean and variance $1/2$. The message passing algorithm we shall develop based on (21) is not optimal, but such separation facilitates approximation and computation, which will be discussed in Section V-B. Since two parts in (21) share the same factor graph, we treat one of them and omit the superscripts:

$$y_\mu = \sqrt{\gamma_s}\sum_{k=1}^{N} s_{\mu k} x_k + w_\mu \tag{22}$$

where $\mu \in 1, 2, \ldots, M$ and $k \in 1, 2, \ldots, N$ index the measurements and the input "symbols," respectively. For simplicity, we ignore the dependence of the symbols $\{X_k\}$ for now, which shall be addressed toward the end of this section. Each X_k then corresponds to a symbol node and each Y_μ corresponds to a measurement node, where the joint distribution of all $\{X_k\}$ and $\{Y_\mu\}$ are decomposed into a product of $M + N$ factors, one corresponding to each node. For every (μ, k), symbol node k and measurement node μ are connected by an edge if $s_{\mu k} \neq 0$. A simple example is shown in Fig. 1 for 5 measurements and 3 neighbors each with 4 messages, i.e., $M = 5$, $K = 3$, and $N = 3 \times 4 = 12$. The actual message chosen by each neighbor for broadcast is marked by a dark node, i.e., $w_1 = 3, w_2 = 2, w_3 = 1$.

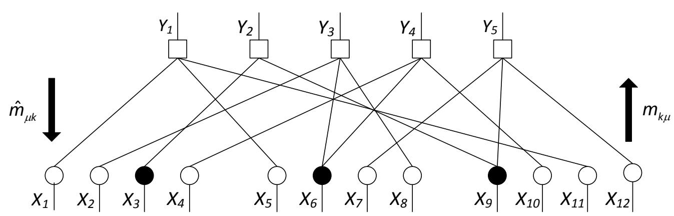

Fig. 1. The Forney-style factor graph of coded mutual broadcast.

B. The Message-Passing Algorithm

In general, an iterative message-passing algorithm involves two steps in each iteration, where a message (or belief, which shall be distinguished from an information message) is first sent from each symbol node to every measurement node it is connected to, and then a new set of messages are computed and sent in the reverse direction, and so forth. The algorithm performs exact inference within finite number of iterations if there are no loops in the graph (the graph becomes a tree if it remains connected), while it provides in general a good approximation for loopy graphs as the one in the current problem.

For convenience, let $\partial\mu$ (resp. ∂k) denote the subset of symbol nodes (resp. measurement nodes) connected directly to measurement node μ (resp. symbol node k), called its neighborhood. [2] $|\partial\mu|$ (resp. $|\partial k|$) represents the cardinality of the neighborhood of measurement node μ (resp. symbol node k). Also, let $\partial\mu\backslash k$ denote the neighborhood of measurement

[2] This is to be distinguished from the notion of neighborhood in the wireless network defined in Section II-C

node μ excluding symbol node k and let $\partial k \backslash \mu$ be similarly defined.

The message-passing algorithm given as Algorithm 1, decodes the information indexes $w_1, \dots, w_K$, and is ready for implementation. The key steps are described as follows. The superscripts $i = 1, 2$ in Algorithm 1 represent the real and imaginary parts, respectively. Here, $\mathsf{E}\left\{ X \,\middle|\, Y = y; \sigma^2 \right\}$ and $\mathsf{var}\left\{ X \,\middle|\, Y = y; \sigma^2 \right\}$ represents the conditional mean and variance of the input given the Gaussian channel output $Y = X + W$ with $W \sim \mathcal{N}(0, \sigma^2)$ is equal to y. Mathematically, assume X has cumulative distribution function $P_X(x)$, then for $i = 1, 2, \dots,$

$$\mathsf{E}\left\{ X^i \,\middle|\, Y = y; \sigma^2 \right\} = \frac{\int x^i e^{-\frac{(y-x)^2}{2\sigma^2}} \mathrm{d}P_X(x)}{\int e^{-\frac{(y-x)^2}{2\sigma^2}} \mathrm{d}P_X(x)}\,, \tag{23}$$

and

$$\mathsf{var}\left\{ X \,\middle|\, Y = y; \sigma^2 \right\} = \mathsf{E}\left\{ X^2 \,\middle|\, Y = y; \sigma^2 \right\} - \left(\mathsf{E}\left\{ X \,\middle|\, Y = y; \sigma^2 \right\} \right)^2, \tag{24}$$

where $\int \cdot \, \mathrm{d}P_X(x)$ in (23) denotes the Riemann-Stieltjes integral.

Algorithm 1 Message-Passing Decoding Algorithm

Input: $\boldsymbol{S}, \boldsymbol{Y}, \gamma_s, M_s, q$.
Initialization:
$z_{\mu k}^{0,i} \leftarrow y_\mu^i / (\sqrt{\gamma_s} s_{\mu k})$ for all $s_{\mu k} \neq 0$ and $i = 1, 2$.
Initialize $\tau^{0,i}$ to a large positive number for $i = 1, 2$.
Main iterations:
for $t = 1$ to $T - 1$ **do**
  for all μ, k with $s_{\mu k} \neq 0$ and $i = 1, 2$ **do**
    $m_{k\mu}^{t,i} \leftarrow \mathsf{E}\left\{ X \,\middle|\, Y = \frac{\sum_{\nu \in \partial k \backslash \mu} z_{\nu k}^{t-1,i}}{|\partial k| - 1}; \frac{\tau^{t-1,i}}{|\partial k| - 1} \right\}$.
    $(\sigma_{k\mu}^{t,i})^2 \leftarrow \mathsf{var}\left\{ X \,\middle|\, Y = \frac{\sum_{\nu \in \partial k \backslash \mu} z_{\nu k}^{t-1,i}}{|\partial k| - 1}; \frac{\tau^{t-1,i}}{|\partial k| - 1} \right\}$.
    $z_{\mu k}^{t,i} \leftarrow \frac{1}{\sqrt{\gamma_s} s_{\mu k}} \left(y_\mu^i - \sqrt{\gamma_s} \sum_{j \in \partial \mu \backslash k} s_{\mu j} m_{j\mu}^{t,i} \right)$.
  end for
  $\tau^{t,i} \leftarrow \frac{1}{\sum_\mu |\partial \mu|} \sum_\mu |\partial \mu| \sum_{j \in \partial \mu} (\sigma_{j\mu}^{t,i})^2 + \frac{1}{2\gamma_s} M_s q (1 - q)$ for $i = 1, 2$.
end for
$m_k^i \leftarrow \mathsf{E}\left\{ X \,\middle|\, Y = \frac{\sum_{\nu \in \partial k} z_{\nu k}^{T-1,i}}{|\partial k| - 1}; \frac{\tau^{T-1,i}}{|\partial k| - 1} \right\}$ for all k, $i = 1, 2$.
Output: $\hat{w}_k = \arg\max_{j = 1, \dots, 2^l} \left| m_{(k-1)2^l + j}^1 + \sqrt{-1}\, m_{(k-1)2^l + j}^2 \right|$, $k = 1, \dots, K$.

In the following, we derive Algorithm 1 starting from (22) which is valid for both real and imaginary parts in (21). It is a simplification of the original iterative BP algorithm, which iteratively computes the marginal *a posteriori* distribution of all symbols given the measurements, assuming that the graph is free of cycles. For each $k \in \partial \mu$ (hence $\mu \in \partial k$), let $\left\{ V_{k\mu}^t(x) \right\}$ represent the message from symbol node k to measurement node μ at the t-th iteration and $\left\{ U_{\mu k}^t(x) \right\}$ represent the message in the reverse direction. Each message is basically the belief (in terms of a probability density (or mass) function) the algorithm has accumulated about the corresponding symbol based on the measurements on the subgraph traversed so far, assuming it is a tree. Let $p_X(x)$ denote the *a priori* probability density function of X. In the t-th iteration, we compute

$$V_{k\mu}^t(x) \propto p_X(x) \prod_{\nu \in \partial k \backslash \mu} U_{\nu k}^{t-1}(x) \tag{25a}$$

for all (k, μ) with $s_{\mu k} \neq 0$, and then

$$U_{\mu k}^t(x) \propto \int_{(x_j)_{\partial \mu \backslash k}} \exp\Bigg[-\Big(y_\mu - \sqrt{\gamma_s} s_{\mu k} x - \sqrt{\gamma_s} \sum_{j \in \partial \mu \backslash k} s_{\mu j} x_j \Big)^2 \Bigg] \left(\prod_{j \in \partial \mu \backslash k} V_{j\mu}^t(x_j) \mathrm{d}x_j \right) \tag{25b}$$

where $\int_{(x_j)_{\partial \mu \backslash k}}$ denotes integral over all x_j with $j \in \partial \mu \backslash k$, and $V(x) \propto u(x)$ means that $V(x)$ is proportional to $u(x)$ with proper normalization such that $\int_{-\infty}^{\infty} V(x) \mathrm{d}x = 1$. In case X is a discrete random variable, the integral shall be replaced by a sum over the alphabet of X. In this problem, X follows a mixture of discrete and continuous distributions, so the expectation can be decomposed as an integral and a sum.

The complexity of computing the integral in (25b) is exponential in $|\partial \mu| = \mathcal{O}(qN)$, which is in general infeasible for the problem at hand. However, as $qN \gg 1$, the computation carried out at each measurement node admits a good approximation by using the central limit theorem. A similar technique has been used in the CDMA detection problem, for fully-connected bipartite graph in [16]–[18], and for a graph with large node degrees in [19].

To streamline (25a) and (25b), we introduce $m_{k\mu}^t$ and $(\sigma_{k\mu}^t)^2$ for all (μ, k) pairs with $s_{\mu k} \neq 0$ to represent the mean and variance of a random variable with distribution $V_{k\mu}^t(x)$. Using Gaussian approximation, one can reduce the message-passing algorithm to iteratively computing the following messages:

$$m_{k\mu}^t = \mathsf{E}\left\{ X \,\middle|\, Y = \frac{\sum_{\nu \in \partial k \backslash \mu} z_{\nu k}^{t-1}}{|\partial k| - 1}; \frac{\tau^{t-1}}{|\partial k| - 1} \right\} \tag{26a}$$

$$(\sigma_{k\mu}^t)^2 = \mathsf{var}\left\{ X \,\middle|\, Y = \frac{\sum_{\nu \in \partial k \backslash \mu} z_{\nu k}^{t-1}}{|\partial k| - 1}; \frac{\tau^{t-1}}{|\partial k| - 1} \right\} \tag{26b}$$

$$z_{\mu k}^t = \frac{1}{\sqrt{\gamma_s} s_{\mu k}} \left(y_\mu - \sqrt{\gamma_s} \sum_{j \in \partial \mu \backslash k} s_{\mu j} m_{j\mu}^t \right) \tag{26c}$$

$$\tau^t = \frac{1}{\sum_\mu |\partial \mu|} \sum_\mu |\partial \mu| \sum_{j \in \partial \mu} (\sigma_{j\mu}^t)^2 + \frac{1}{2\gamma_s} M_s q (1 - q) \tag{26d}$$

where (26a) and (26b) calculate the conditional expectation and variance, respectively. The detailed derivation is relegated

to Appendix C. At the T-th iteration, the approximated posterior mean of x_k can be expressed as

$$m_k = \mathsf{E}\left\{ X \,\middle|\, Y = \frac{\sum_{\nu \in \partial k} z_{\nu k}^{T-1}}{|\partial k| - 1}; \frac{\tau^{T-1}}{|\partial k| - 1} \right\}. \quad (27)$$

It is time consuming to compute (26a) and (26b) for all (μ, k) pairs with $s_{\mu k} \neq 0$, especially in the case of large matrix $\boldsymbol{S}$. We can use the following two approximation techniques to further decrease the computational complexity. First, $|\partial k|$ in (26a), (26b) and (27) is replaced by its mean value $M_s q(1-q)$. Second, we use interpolation and extrapolation to further reduce the computation complexity of (26a), (26b) and (27). Specifically, in each iteration t, we only compute the conditional mean and variance for some chosen y's, i.e., we choose $y_1^t < y_2^t < \cdots < y_n^t$ which is a partition of an interval depending on τ^{t-1}, compute

$$a_j^t = \mathsf{E}\left\{ X \,\middle|\, Y = y_j^t; \frac{\tau^{t-1}}{M_s q(1-q) - 1} \right\}, \quad (28)$$

$$b_j^t = \mathsf{var}\left\{ X \,\middle|\, Y = y_j^t; \frac{\tau^{t-1}}{M_s q(1-q) - 1} \right\} \quad (29)$$

for $j = 1, 2, \ldots, n$, and then use those values to calculate (26a), (26b) by interpolation or extrapolation. To be more precise, for any pair μ, k with $s_{\mu k} \neq 0$, suppose y_j^t and y_{j+1}^t are chosen to be the closest to $y = \frac{\sum_{\nu \in \partial k \setminus \mu} z_{\nu k}^{t-1, i}}{|\partial k| - 1}$, then $m_{k\mu}^t$ and $(\sigma_{k\mu}^t)^2$ can be approximated by

$$m_{k\mu}^t = a_j^t + \frac{y - y_j^t}{y_{j+1}^t - y_j^t}(a_{j+1}^t - a_j^t), \quad (30)$$

$$(\sigma_{k\mu}^t)^2 = b_j^t + \frac{y - y_j^t}{y_{j+1}^t - y_j^t}(b_{j+1}^t - b_j^t). \quad (31)$$

We now revisit the assumption that $\boldsymbol{X}$ has independent elements. In fact, $\boldsymbol{X}$ consists of K sub-vectors of length 2^l, where the entries of each sub-vector are all zero except for one position corresponding to the transmitted message. After obtaining the approximated posterior mean $\widetilde{m}_k$ by incorporating both real and imaginary parts calculated from (27), Algorithm 1 outputs the position of the element with the largest magnitude in each of the K sub-vectors of $[\widetilde{m}_1, \ldots, \widetilde{m}_N]$. In fact the factor graph Fig. 1 can be modified to include K additional nodes, each of which puts a constraint on one sub-vector. Slight improvement over Algorithm 1 can be obtained by carrying out message passing on the modified graph.

VI. Numerical Results

In order for a fair comparison, we assume the same power constraint for both the sparse recovery scheme and random-access schemes, i.e., the average transmit power in each active slot (in which the node transmits energy) is the same. Recall that, in the proposed scheme, where the frame length is M_s, the SNR in the model (18) is $\gamma_s = \gamma M_s q(1-q)$. We choose the same transmission probability in each slot for sparse recovery scheme and slotted ALOHA, i.e., $q = p = 1/(c+1)$. Also, the transmission probability in each slot for CSMA is $(1 - e^{-c})/c$ (see Appendix III-B), which is close to $1/(c+1)$ when c is large. The three schemes consume approximately the same amount of average power over any period of time.

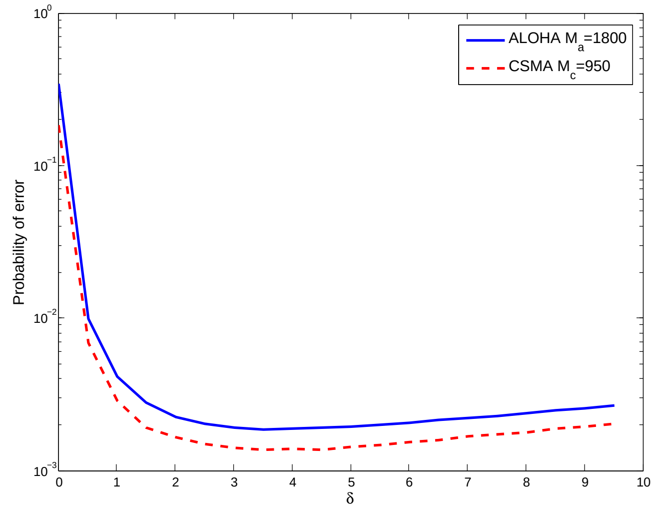

Fig. 2. Low bounds for error probability in slotted-ALOHA and CSMA for different threshold δ in the case of $l = 10$.

Without loss of generality, let one unit of distance be 1 meter. Consider a wireless network of 1000 nodes uniformly distributed in a square with side length 500 meters. The nodes form a Poisson point process in the square conditioned on the node population. Suppose the path-loss exponent $\alpha = 4$. The threshold of channel gain to define neighborhood is set to $\theta = 10^{-6}$. It means that if the transmit power for a node one meter away is 60 dB, then the SNR attenuates to 0 dB at $= 10^{6/\alpha} \approx 31$ meters in the absence of fading, i.e., the coverage of the neighborhood of a node is typically a circle of radius 31 meters. According to (8), a node near the center of the square (without boundary effect) has on average $c \approx 11$ neighbors.

We consider two cases for the length of broadcast message $l = 5$ and 10 bits. In random access schemes, a packet of L bits consists of l-bit message and $\lceil \log_2 c \rceil$ additional bits to identify the sender. Fig. 2 shows that $T = 3.5$ minimizes the lower bounds for $\mathsf{P}_\mathrm{a}^\mathrm{e}$ in (12) and $\mathsf{P}_\mathrm{c}^\mathrm{e}$ in (15) in the case of $l = 10$.

The metric for performance comparison is the probability for one node to miss one specific neighbor, averaged over all pairs of neighboring nodes in the network. Suppose the transmit power of each node is $\gamma = 60$ dB. First consider one realization of the network where each node has $c \approx 11$ neighbors on average and $l = 5$ bits to broadcast, so that on average $cl \approx 55$ bits are to be collected by each node. In slotted ALOHA, at least 4 additional bits are needed to identify a sender out of 1000 nodes, so we let $L = 9$. In Fig. 3, the error performance of slotted ALOHA and CSMA for $\delta = 0.5$ is compared with that of the sparse recovery scheme with the message-passing algorithm. The simulation result shows the sparse recovery scheme significantly outperforms slotted ALOHA and CSMA, even compared with the minimum of the lower bounds computed from (12) and (15) for $\delta = 3.5$. For example, to achieve 1% error rate, the sparse recovery scheme takes fewer than 300 symbols. Slotted ALOHA and

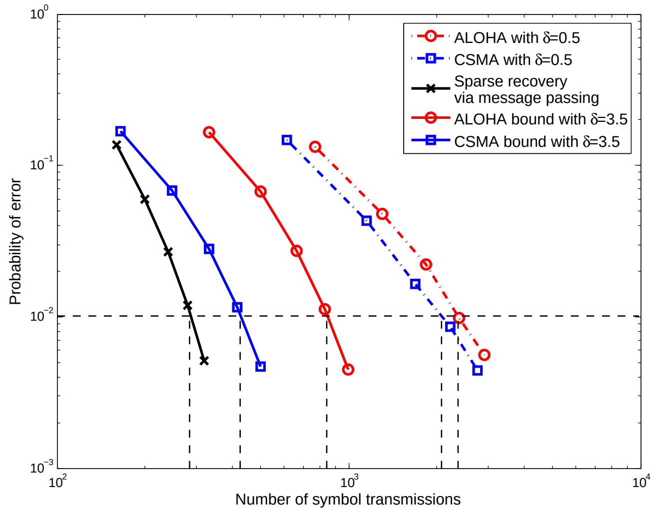

Fig. 3. Performance comparison between sparse recovery and random access. Each node transmits a 5-bit message.

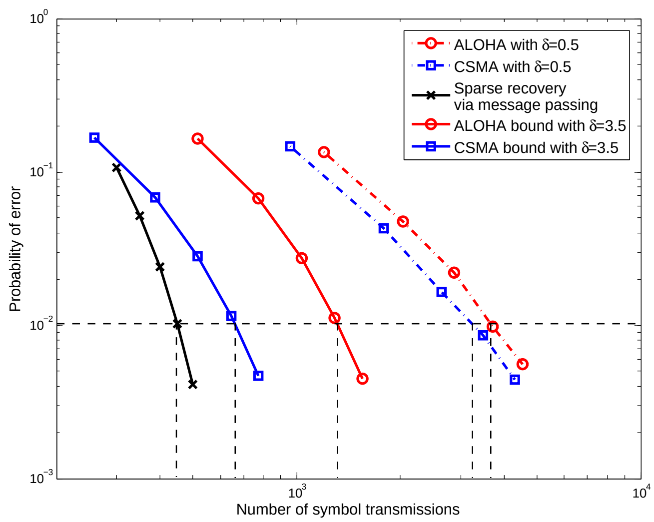

Fig. 4. Performance comparison between sparse recovery and random access. Each node transmits a 10-bit message.

CSMA take no less than 800 and 400 symbols according to the bounds in (12) and (15), respectively. In fact, slotted ALOHA and CSMA with threshold $\delta = 0.5$ take more than 2000 symbols. Similar comparison is observed for several other SINR thresholds δ around 0.5 and the performance of ALOHA and CSMA are not good for $\delta \geq 1$ because the messages from weaker neighbors may never be successfully delivered. Some additional supporting numerical evidence is, however, omitted due to space limitations.

Fig. 4 repeats the experiment of Fig. 3 with 10-bit messages. The sparse recovery scheme has significant gain compared with slotted ALOHA and CSMA. For example, to achieve the error rate of 1%, the sparse recovery scheme takes about 450 symbols, whereas slotted ALOHA and CSMA take at least 1000 and 650 symbols, respectively.

In Fig. 5, we simulate the same network with different nominal SNRs, i.e., γ varies from 50 dB to 70 dB. In the case that each node transmits a 5-bit message, the frame length is

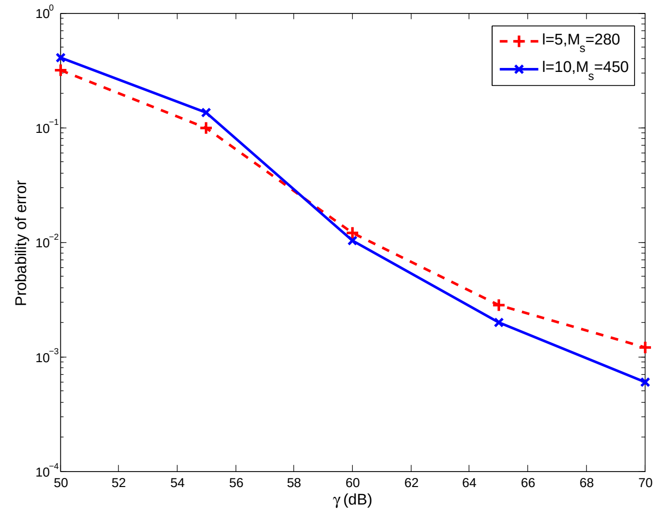

Fig. 5. Performance of sparse recovery scheme in different nominal SNR (γ).

chosen to be 280 symbols. It can be seen from the figure that the probability of error decreases with the increase of SNR. The performance is similar when each node transmits a 10-bit message and the frame consists of 450 symbols.

VII. Concluding Remarks

The idea of using on-off signalling to achieve full-duplex communication using half-duplex radios applies to general peer-to-peer networks, and is not limited to mutual broadcast traffic focused on in this paper. The sparse recovery scheme with random signatures are of course most suitable for the situation where the broadcast messages consist of a relatively small number of bits. If each node has many bits to send, a structured code with low decoding complexity is needed for the scheme to be practical. This is ongoing research.

The proposed scheme can also serve as a highly desirable sub-layer of any network protocol stack to provide the important function of simultaneous message exchange among peers. This sub-layer provides the missing link in many advanced resource allocation schemes, where it is often *assumed* that nodes are provided the state and/or demand of their peers.

Appendix A
Proof of Proposition 1

Let $\hat{\Phi}_1^p$ be an independent thinning of $\hat{\Phi}_1$ with retention probability p to represent the transmitting nodes. It is easy to see that $\hat{\Phi}_1^p$ is an independent marked p.p.p. with intensity λp. Denote

$$I = \sum_{(X_i, \mathcal{G}_i) \in \hat{\Phi}_1^p} G_{0i} |X_{0i}|^{-\alpha}, \tag{32}$$

then we have

$$\begin{aligned} \mathsf{E}\left\{ \mathsf{P}_{\mathrm{a}}^{\mathrm{s}}(X, \mathcal{G}, \hat{\Phi}_1) \right\} &= \mathsf{E}\left\{ \mathsf{E}\left\{ \mathbf{1}\left(\frac{\gamma G |X|^{-\alpha}}{\gamma I + 1} \geq \delta \right) \middle| \hat{\Phi}_1^p \right\} \right\} \\ &= \mathsf{E}\left\{ \mathsf{P}\left\{ G|X|^{-\alpha} \geq \delta \left(I + \frac{1}{\gamma} \right) \middle| \hat{\Phi}_1^p \right\} \right\} \end{aligned}$$

$$= \mathsf{E}\left\{\left(\frac{\theta}{\delta}\right)^{\frac{2}{\alpha}}\left(I+\frac{1}{\gamma}\right)^{-\frac{2}{\alpha}}\right\} \tag{33}$$

$$= \left(\frac{\theta}{\delta}\right)^{\frac{2}{\alpha}}\int_{-\infty}^{\infty}\left|i+\frac{1}{\gamma}\right|^{-\frac{2}{\alpha}} p_I(i)\mathrm{d}i \tag{34}$$

where (33) is derived from (5) and p_I is the pdf of interference I.

According to [12], the Laplace transform of p_I can be expressed as

$$\mathcal{L}_{p_I}(s) = \exp\left\{-\lambda p s^{2/\alpha}\frac{2\pi^2}{\alpha\sin(2\pi/\alpha)}\right\}. \tag{35}$$

Therefore, the Fourier transform[3] of p_I can be obtained by replacing s in (35) by $\iota\omega$ with $\iota=\sqrt{-1}$ as

$$\mathcal{F}_{p_I}(\omega) = \exp\left\{-\lambda p(\iota\omega)^{2/\alpha}\frac{2\pi^2}{\alpha\sin(2\pi/\alpha)}\right\}. \tag{36}$$

Since the Fourier transform of $|x|^a$ for $-1<a<0$ is

$$\mathcal{F}_{|x|^a}(\omega) = -\frac{2\sin(a\pi/2)\Gamma(a+1)}{|\omega|^{a+1}}, \tag{37}$$

the Fourier transform of

$$q_I(i) = |i+\frac{1}{\gamma}|^{-\frac{2}{\alpha}} \tag{38}$$

for $\alpha>2$ can be expressed as

$$\mathcal{F}_{q_I}(\omega) = e^{\iota\omega/\gamma}\frac{2\sin(\pi/\alpha)\Gamma(1-2/\alpha)}{|\omega|^{1-2/\alpha}}. \tag{39}$$

Since the integral in (34) can be viewed as the Fourier transform of $p_I(i)q_I(i)$ at $\omega=0$, it can be calculated as the convolution of $\mathcal{F}_{p_I}(\omega)$ and $\mathcal{I}_{q_I}(\omega)$ at $\omega=0$ [20]. Therefore, by (36) and (39), we have

$$\int_{-\infty}^{\infty}\left|i+\frac{1}{\gamma}\right|^{-\frac{2}{\alpha}} p_I(i)\mathrm{d}i = \frac{1}{2\pi}\mathcal{F}_{p_I}(\omega)*\mathcal{F}_{q_I}(\omega)\Big|_{\omega=0} \tag{40}$$

where $*$ is the convolution operator. Therefore, according to (11), (34) and (40), the error probability $\mathsf{P}_{\mathrm{a}}^{\mathrm{e}}$ can be lower bounded as

$$\mathsf{P}_{\mathrm{a}}^{\mathrm{e}} \geq \left(\max\left\{0, 1-\frac{1}{2\pi}p(1-p)\left(\frac{\theta}{T}\right)^{\frac{2}{\alpha}} \mathcal{F}_{p_I}(\omega)*\mathcal{F}_{q_I}(\omega)\Big|_{\omega=0}\right\}\right)^{M_f}. \tag{41}$$

According to (9), the number of frames in a period of M_a symbol intervals should satisfy

$$M_f \geq M_a\log_2(1+\delta)/L. \tag{42}$$

Therefore, (12) in Proposition 1 follows by combining (41) and (42).

[3]The reasons to work with Fourier transform in lieu of Laplace transform are: 1) The inverse Fourier transform here is easier to calculate; 2) the Fourier transform of $|i+\frac{1}{\gamma}|^{-\frac{2}{\alpha}}$ has a closed form.

Appendix B
Proof of Proposition 2

Denote G_i as the fading coefficients node $X_i \in \hat{\Phi}_1$ to node X_0. Define the following indicators for node X

$$F_1 = \mathbf{1}(T_0 > T) \tag{43}$$

$$F_2 = \mathbf{1}\left(T_i > T,\ \forall X_i \in \hat{\Phi}_1 \text{ with } G_i|X_i - X|^{-\alpha} \geq \theta\right) \tag{44}$$

$$F_3 = \mathbf{1}\left(\gamma G|X|^{-\alpha} \geq \delta\right) \tag{45}$$

where $F_1=1$ if and only if the timer of X expires before that of X_0, $F_2=1$ if and only if X's timer expires sooner than those of all its neighbors excluding X_0, $F_3=1$ if and only if the received SNR from node X to node X_0 exceeds the threshold δ. In order for the transmission to be successful, we must have $F_1=F_2=F_3=1$. That is

$$\mathsf{E}\left\{\mathsf{P}_{\mathrm{c}}^{\mathrm{s}}(X,\mathcal{G},\hat{\Phi}_1)\right\} \leq \mathsf{E}\left\{F_1F_2F_3\right\}. \tag{46}$$

Conditioned on $T=\varsigma$, we express the indicator $e^{(2)}$ as the value of some extremal shot-noise [12, Section 2.4]. For fixed ς, define the indicator of the event that X_i is a neighbor of X and it has a timer smaller than ς:

$$L(X, X_i, G_i, T_i) = \mathbf{1}\left(G_i|X_i - X|^{-\alpha} \geq \theta \text{ and } T_i < \varsigma\right) \tag{47}$$

for all $(X_i, \mathcal{G}_i, T_i) \in \hat{\Phi}_1$. Define the extremal shot-noise at node X as

$$Z_{\hat{\Phi}_1}(X) = \max_{(X_i,\mathcal{G}_i,T_i)\in\hat{\Phi}_1} L(X, X_i, G_i, T_i). \tag{48}$$

Note that $Z_{\hat{\Phi}_1}(X)$ takes only two values 0 or 1 and consequently

$$\mathsf{E}\left\{F_2\Big|T=\varsigma\right\} = \mathsf{P}\left\{Z_{\hat{\Phi}_1}(X) \leq 0\Big|T=\varsigma\right\}. \tag{49}$$

By [12, Proposition 2.4.2], (49) can be further calculated as

$$\begin{aligned}
&\mathsf{E}\left\{F_2\Big|T=\varsigma\right\}\\
&= \exp\left\{-\lambda\int_{\mathbb{R}^2}\int_0^\infty\int_0^1 \mathbf{1}\left(L(X,x,g,t)=1\right)e^{-g}\mathrm{d}t\mathrm{d}g\mathrm{d}x\right\}\\
&= \exp\left\{-2\pi\lambda\varsigma\int_0^\infty\int_0^\infty \mathbf{1}\left(gr^{-\alpha}\geq\theta\right)re^{-g}\mathrm{d}g\mathrm{d}r\right\}\\
&= e^{-c\varsigma}
\end{aligned} \tag{50}$$

where c is the average number of neighbors defined in (8).

Therefore, according to (46), we have

$$\mathsf{E}\left\{\mathsf{P}_{\mathrm{c}}^{\mathrm{s}}(X,\mathcal{G},\hat{\Phi}_1)\right\} \leq \left(\frac{\theta\gamma}{\delta}\right)^{\frac{2}{\alpha}}\int_0^1(1-\varsigma)e^{-c\varsigma}\mathrm{d}\varsigma \tag{51}$$

$$= \frac{1}{c^2}\left(\frac{\theta\gamma}{\delta}\right)^{\frac{2}{\alpha}}\left(e^{-c}+c-1\right) \tag{52}$$

where (51) is derived from the the uniform distribution of T_0, (5) and (49).

According to (9), the number of frames in a period of M_c symbol intervals should satisfy

$$M_f \geq M_c\log_2(1+\delta)/L. \tag{53}$$

Therefore, Proposition 2 follows by combining (52) and (53).

As a by-product, by averaging over ς in (50), which is uniformly distributed on $[0,1]$, the probability that a given node captures the channel to transmit in each slot can be calculated as $(1-e^{-c})/c$.

APPENDIX C
DERIVATION OF MESSAGE COMPUTATION (26)

We derive (26) from (25). Denote $\Delta_{\mu k} = \sum_{j\in\partial\mu\setminus k} s_{\mu j} x_j$. The key to the simplification is to recognize that $\Delta_{\mu k}$ is approximately Gaussian. To be precise, if $\{x_j\}_{j\in\partial\mu\setminus k}$ were independent (conditioned on the observations traversed so far on the graph), then, by central limit theorem, $\Delta_{\mu k}$ converges weakly to a Gaussian random variable, whose mean is

$$v_{\mu k}^t = \sum_{j\in\partial\mu\setminus k} s_{\mu j} m_{j\mu}^t \tag{54}$$

and variance is

$$(\sigma_{\mu k}^t)^2 = \sum_{j\in\partial\mu\setminus k} s_{\mu j}^2 (\sigma_{j\mu}^t)^2 \,. \tag{55}$$

Using the preceding Gaussian approximation, (25b) can be calculated by a change of probability measure as

$$\begin{aligned} U_{\mu k}^t(x) \propto & \int_{-\infty}^{\infty} \exp\left[-\left(y_\mu - \sqrt{\gamma_s} s_{\mu k} x - \sqrt{\gamma_s}\Delta\right)^2\right] \\ & \cdot \frac{1}{\sqrt{2\pi(\sigma_{\mu k}^t)^2}} \exp\left[-\frac{1}{2(\sigma_{\mu k}^t)^2}\left(\Delta - v_{\mu k}\right)^2\right] \mathrm{d}\Delta \\ \propto & \exp\left[-\frac{1}{2\tau_{\mu k}^t}(x - z_{\mu k}^t)^2\right] \end{aligned} \tag{56}$$

where $z_{\mu k}^t$ is defined in (26c) and

$$\tau_{\mu k}^t = \sum_{j\in\partial\mu\setminus k} (\sigma_{j\mu}^t)^2 + \frac{1}{2\gamma_s s_{\mu k}^2} \,. \tag{57}$$

Using law of large numbers, we further approximate $\tau_{\mu k}^t$ by its average over all (μ,k) pairs with $s_{\mu k}\neq 0$, i.e., $\tau_{\mu k}^t$ is replaced by

$$\begin{aligned} \tau^t &= \frac{1}{\sum_\mu |\partial\mu|} \sum_k \sum_{\mu\in k} \sum_{j\in\partial\mu\setminus k} (\sigma_{j\mu}^t)^2 + \frac{1}{2\gamma_s s_{\mu k}^2} \\ &\approx \frac{1}{\sum_\mu |\partial\mu|} \sum_\mu |\partial\mu| \sum_{j\in\partial\mu} (\sigma_{j\mu}^t)^2 + \frac{1}{2\gamma_s} M_s q(1-q) \end{aligned} \tag{58}$$

as shown in (26d).

Now we have $U_{\mu k}^t \sim \mathcal{N}(z_{\mu k}^t, \tau^t)$, so it is easy to see that

$$\prod_{\nu\in\partial k\setminus\mu} U_{\nu k}^t \sim \mathcal{N}\left(\frac{\sum_{\nu\in\partial k\setminus\mu} z_{\nu k}^t}{|\partial k|-1}, \frac{\tau^t}{|\partial k|-1}\right) \,. \tag{59}$$

According to (25a), $V_{k\mu}^{t+1}(x)$ can be viewed as the conditional distribution $p_{X|Y}\left(x \,\middle|\, \frac{\sum_{\nu\in\partial k\setminus\mu} z_{\nu k}^t}{|\partial k|-1}\right)$, where Y is the output of a Gaussian channel with noise $W \sim \mathcal{N}\left(0, \frac{\tau^t}{|\partial k|-1}\right)$. Therefore, by definition, $m_{k\mu}^{t+1}$ and $(\sigma_{k\mu}^{t+1})^2$ can be expressed as the conditional mean and conditional variance as in (26a) and (26b).

REFERENCES

[1] L. Zhang and D. Guo, “Wireless peer-to-peer mutual broadcast via sparse recovery,” in *Proc. IEEE Int. Symp. Inform. Theory*, pp. 1901–1905, July-Aug. 2011.

[2] D. Guo and L. Zhang, “Rapid on-off-division duplex for mobile ad hoc neworks,” in *Proc. Allerton Conf. Commun., Control, & Computing*, Monticello, IL, USA, 2010.

[3] I. D. Schizas, A. Ribeiro, and G. B. Giannakis, “Consensus in ad hoc WSNs with noisy links-part I: Distributed estimation of deterministic signals,” *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 350–364, 2008.

[4] I. D. Schizas, G. B. Giannakis, S. I. Roumeliotis, and A. Ribeiro, “Consensus in ad hoc WSNs with noisy links-part II: Distributed estimation and smoothing of random signals,” *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1650–1666, 2008.

[5] O. Simeone, U. Spagnolini, Y. Bar-Ness, and S. Strogatz, “Distributed synchronization in wireless networks,” *IEEE Signal Processing Mag.*, vol. 25, pp. 81–97, Sep 2008.

[6] D. Needell and J. A. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” *Applied and Computational Harmonic Analysis*, vol. 26, pp. 301–321, 2009.

[7] D. L. Donoho, A. Maleki, and A. Montanari, “Message passing algorithms for compressed sensing: I. motivation and construction and II. analysis and validation,” in *Proc. IEEE Inform. Theory Workshop*, Cairo, Egypt, Jan. 2010.

[8] A. Maleki, *Approximate message passing algorithms for compressed sensing*. PhD thesis, Standord University, 2011.

[9] W. Dai and O. Milenkovic, “Subspace pursuit for compressive sensing signal reconstruction,” *IEEE Trans. Inform. Theory*, vol. 55, pp. 2230–2249, 2009.

[10] D. Baron, S. Sarvotham, and R. G. Baraniuk, “Bayesian compressive sensing via belief propagation,” *IEEE Trans. Signal Process.*, vol. 58, pp. 269–280, 2010.

[11] L. Applebaum, W. U. Bajwa, M. F. Duarte, and R. Calderbank, “Asynchronous code-division random access using convex optimization,” *Physical Communication*, vol. 5, no. 2, pp. 129 – 147, 2012.

[12] F. Baccelli and B. Błaszczyszyn, *Stochastic Geometry and Wireless Networks: Volume I Theory and Volume II Applications*, vol. 4 of *Foundations and Trends in Networking*. NoW Publishers, 2009.

[13] S. A. Borbash, A. Ephremides, and M. J. McGlynn, “An asynchronous neighbor discovery algorithm for wireless sensor networks,” in *Ad Hoc Networks*, vol. 5, pp. 998–1016, Sep. 2007.

[14] J. Luo and D. Guo, “Neighbor discovery in wireless ad hoc networks based on group testing,” in *Proc. Allerton Conf. Commun., Control, & Computing*, Monticello, IL, USA, 2008.

[15] L. Zhang, J. Luo, and D. Guo, “Neighbor discovery for wireless networks via compressed sensing,” *Performance Evaluation*, 2012.

[16] Y. Kabashima, “A CDMA multiuser detection algorithm on the basis of belief propagation,” *Journal of Physics A: Mathematical and General*, vol. 26, pp. 11111–11121, Oct. 2003.

[17] T. Tanaka and M. Okada, “Approximate belief propagation, density evolution, and statistical neurodynamics for CDMA multiuser detection,” *IEEE Trans. Inform. Theory*, vol. 51, pp. 700–706, Feb. 2005.

[18] P. H. Tan and L. Rasmussen, “Belief propagation for coded multiuser detection,” in *Proc. IEEE Int. Symp. Inform. Theory*, pp. 1919 –1923, July 2006.

[19] D. Guo and C.-C. Wang, “Multiuser detection of sparsely spread CDMA,” *IEEE J. Select. Areas Commun.*, vol. 26, pp. 421–431, Apr. 2008.

[20] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals & systems (2nd ed.)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996.